

Longitudinal Analysis of Categorical Epidemiological Data: A Study of Three Mile Island

by Stephen E. Fienberg,* Evelyn J. Bromet,[†] Dean Follmann,[‡] Diane Lambert,* and Sherryl M. May**

The accident at the Three Mile Island nuclear power plant in 1979 led to an unprecedented set of events with potentially life threatening implications. This paper focusses on the analysis of a longitudinal study of the psychological well-being of the mothers of young children living within 10 miles of the plant. The initial analyses of the data utilize loglinear/logit model techniques from the contingency table literature, and involve the fitting of a sequence of logit models. The inadequacies of these analyses are noted, and a new class of mixture models for logistic response structures is introduced to overcome the noted shortcomings. The paper includes a brief outline of the methodology relevant for the fitting of these models using the method of maximum likelihood, and then the model is applied to the TMI data. The paper concludes with a discussion of some of the substantive implications of the mixture model analysis.

Introduction

The analysis of epidemiological data has evolved over the past 25 years with the development of new methodology for the analysis of categorical data using log-linear and logit models. This methodology has led to new formulations for epidemiological questions, and these in turn have stimulated additional methodological developments for the analysis of categorical data. In this paper we review some of the methodology for the analysis of categorical data from the special perspective of longitudinal epidemiological data, and we illustrate the methodology in the context of a study of the mental health effects of the Three Mile Island nuclear accident, which occurred in the spring of 1979.

Implicit in the study of longitudinal data is the recognition that we are interested in the study of change—in the context of the Three Mile Island study there are both changes in attitudes and changes in mental health status. The classical cross-sectional study is useful for learning about populations as they are; to study change we need observations or information about two or more periods of time. Studies often have looked at *net change* by comparing the results of two or more separate cross sections, but epidemiological researchers have long

noted that, to study change at a level that will allow for understanding of the processes that cause or are related to change, we really need to follow individual units longitudinally. The primary longitudinal design used in epidemiology is the cohort study, in which groups of individuals are assembled and classified with regard to exposures of interest and are followed forward in time to determine the development of disease. Breslow (1) provides an excellent review of cohort analysis in epidemiology for the study of disease incidence or mortality rates. The problem we choose to study in this paper is a specialized version of the broad class considered by Breslow, in which we observe the repeated measurement of a selected set of variables.

Loglinear Model Developments for Epidemiological Data

The historical development of methods for the analysis of categorical data has been described elsewhere (2-4). The brief review here focusses on developments tied in part to epidemiological research.

Although Yule (5) presented the crossproduct ratio as a measure of association in 2×2 contingency tables and developed ideas on association in 2^n tables, it was Bartlett (6) who developed these ideas through his definition of the concept of no second-order interaction in $2 \times 2 \times 2$ tables. Bartlett's model was the forerunner to the logit or logistic models used in this paper. Dyke and Patterson (7) used a generalization of Bartlett's approach in developing a logit model for the knowledge of

*Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA 15213.

[†]Department of Psychiatry and Epidemiology, University of Pittsburgh, Pittsburgh, PA 15213.

[‡]Center for Naval Analysis, Alexandria, VA 22311.

**Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15213.

cancer facts as a function of categorical explanatory variables involving exposure to media. Other key figures in the development of these methods during the 1950's include Berkson, Cochran, Cox, Haldane, and Woolf.

Two major epidemiological studies in the 1960's gave a renewed stimulus to the development of methods for categorical data analysis: the Framingham Heart Study (8) of atherosclerotic disease, and the National Halothane Study (9) of massive hepatic damage subsequent to surgical anesthesia. The Framingham study led to the work of Truett et al. (10) and Walker and Duncan's (11) linear logistic model for polytomous response variables. While working on the halothane study, Bishop (12) rediscovered an iterative procedure proposed for a related categorical data problem by Deming and Stephan (13) and showed how it could be used to solve the likelihood equations associated with the class of loglinear and logit models described by Birch (14). This work led to developments by Goodman, Haberman, Mantel, Plackett, and others.

In the 1970's these two streams of research generated by the Framingham and Halothane studies were joined (15), and a general approach to the analysis of multivariate categorical data emerged. It is this approach to the analysis of longitudinal epidemiological data which we adopt and extend in this paper.

Outline of Paper

In the next section, we present a description of the mental health study of the Three Mile Island (TMI) accident. Then, we illustrate the application of the log-linear/logit model techniques for the analysis of contingency tables in a detailed analysis of the TMI data using a sequence of logit models. A residual item at the end of that section is how to present a succinct summing of the salient features of our analysis.

We next introduce a new class of mixture models for logistic response structures, and we present a brief outline of the methodology relevant for the fitting of these models. Finally, we reanalyze the TMI data using the mixture model approach.

The Three Mile Island Study

The accident at the Three Mile Island (TMI) nuclear power plant in central Pennsylvania began on March 28, 1979 and led to an unprecedented set of events with potentially life threatening implications. Because of the possible grave danger to the surrounding communities, particularly to children who were presumed to be highly susceptible to the hazards of radiation, mothers of young children were ordered to evacuate from the five-mile radius around the plant. In fact, most mothers living within ten miles of the plant evacuated with their children. Since then, controversies surrounding the clean-up operations of the damaged unit and the reopening of the undamaged reactor, and government probes into mismanagement at the plant before and after the accident have ensured continued public awareness of the

Table 1. TMI mother sample.

Time of interview	Sample size
Winter, 1979	327
Spring, 1980	311 (95%)
Fall, 1981	281 (86%)
Fall, 1982	268 (82%)

problems at TMI. As a result, the psychological well-being of residents of the TMI area has been a major public health concern.

In order to evaluate systematically the impact of the stresses associated with the accident, a longitudinal epidemiologic study (16) was designed which focussed on the mental health of mothers of young children living within ten miles of the plant. Mothers were selected because of (a) possible additional stress they experienced as a result of the evacuation order; (b) their concerns about the potential effects of radiation exposure on their children; and (c) data from Great Britain indicating greater vulnerability to depression among mothers of young children (17).

Four waves of interviews were conducted to ascertain changes in psychological symptomatology over time and effects of TMI-related factors. In this set of analyses, the two risk factors of concern are distance of residence from the plant and perceptions of the dangerousness of TMI. With respect to distance, public policy attention had focussed on the five mile radius as being the area at greatest risk for radiation exposure. As a result, the psychological stress of living nearer to the plant was expected to be higher among residents in that area. Preliminary analysis of data collected during the year after the accident indicated that mothers within the five mile radius were in fact experiencing greater symptomatology than mothers living 6–10 miles away. Thus distance was dichotomized for most of the present analyses into 0–5 miles vs. 6–10 miles. With respect to danger, half of the mothers interviewed perceived TMI as dangerous at each interview. This perception can be conceptualized as a risk factor since it might indicate a sense of vigilance and concern about the plant.

Sample

The target sample was assembled from newspaper listings of birth announcements appearing in the three area newspapers for the period January 1, 1978–March 28, 1979. This strategy was selected because of a state law prohibiting access to birth certificates compiled by the state health department. As Table 1 shows, a total of 268 mothers of young children were interviewed on each of four occasions, i.e., 9 months, 12 months, 30 months, and 42 months after the accident. The greatest loss to follow-up occurred during the 18 months between the second and third wave, which was also the longest interval between interview points. Most mothers lost to follow-up had moved out of the area and therefore were ineligible for continued participation in the study. The State Health Department conducted a study of mo-

bility into and out of the TMI area before and after the accident and found no evidence to indicate that the rate increased after the accident. No differences on basic demographic or mental health variables were found between participants and nonparticipants at any wave when the data were examined cross-sectionally.

Demographically, the sample was relatively homogeneous. At the start of the study, the median age was 27–28 years, two-thirds of the mothers being less than 30 years of age. Most of the women had one or two children. About half of the mothers had some formal education beyond high school. Two-thirds of the women had grown up in the same area in which they were currently living, were Protestant, and were not working outside the home. Of the mothers 98% were married and Caucasian. They were also predominantly middle income, having husbands who were primarily engaged in white collar occupations.

Procedure

The interviews were conducted in participants' homes and lasted approximately 90 minutes at each wave. The interviewers were mental health professionals with at least a masters degree in a related field and a minimum of five years of work experience. Although initially respondents were randomly assigned to interviewers, the same interviewer was retained where possible for each follow-up interview in order to reduce attrition.

Mental Health

The present analysis focusses on changes in the TMI risk factors noted above in relation to changes in symptomatology. Current symptom levels were determined from responses to a 90-item self-report checklist (18) administered at each wave. This checklist inquires about symptoms during the two-weeks preceding the interview. The items are rated on a five-point 0-4 distress scale (not at all, a little bit, moderately, quite a bit, extremely). Items cover symptoms related to depression, anxiety, psychosomatic complaints, paranoia, psychotic feelings, and the like. The psychometric properties have been well documented, and normative data are available on a community sample of women. We use as a summary measure in the present analysis the Global Severity Index (GSI) which is the average distress level for the 90 items.

The distribution of GSI scores at each wave was severely skewed, with most mothers exhibiting relatively low scores, e.g., in the first wave, 221 mothers had a score less than 0.5, 85 had a score between 0.5 and 1.10, and 21 had a score greater than 1.10. We attempted to analyze the data for longitudinal purposes in the logarithmic scale and by breaking the distribution into categories. The analyses discussed in the following sections uses several different choices for the GSI categories.

For the moment, suppose that we have trichotomized GSI scores. Thus the basic data of interest consist of the GSI scores and perception of danger at each of four

waves plus distance and they form a $(3 \times 2)^4 \times 2$ contingency table. We have 257 complete observations to spread over these 2592 cells, so that the table is extremely sparse. Even if we had dichotomized the GSI scores there would be 512 cells, and we would still need to find a more succinct way to summarize the data, and to describe its salient features.

Some Approaches from Loglinear Model Theory

We assume that the reader has some familiarity with the basic ideas of loglinear models and their use in the analysis of two-dimensional and three-dimensional contingency tables such as can be found in Fienberg (2,3) or in Grizzle and Koch (19). Throughout this paper we consider the estimation of parameters in such models using the method of maximum likelihood, and we use likelihood ratio tests to check the goodness-of-fit of our models. We focus our attention here on models of the general linear logistic form:

$$\log \left(\frac{p}{1-p} \right) = \sum_i \beta_i x_i \quad (1)$$

where p is the probability of a positive outcome on a dichotomous response variable, and the $\{x_i\}$ are observed values corresponding to a set of explanatory variables. In much of our discussion, the explanatory variables are categorical and the corresponding x_i are indicator variables. The form of expression (1) is sufficiently general to allow for polynomial structures and interactions of the form encountered in the usually linear and loglinear model literature.

For concreteness, we consider a situation where GSI scores have been trichotomized (e.g., high = 3, medium = 2, and low = 1), and we are interested in predicting GSI scores for wave 4 as a function of the scores at waves 1, 2, and 3, and distance from TMI (0-5 miles = 1, 6-10 miles = 2). Thus the data of interest from a $3 \times 3 \times 3 \times 3 \times 2$ table of observed counts, $\{x_{ijkhd}\}$, where i indexes GSI scores at wave 1, j indexes GSI scores at wave 2, k indexes GSI scores at wave 3, h indexes GSI scores at wave 4, d indexes distance from TMI. The counts are given here in Table 2. The most striking feature of Table 2 is that most of the mothers are concentrated in a few cells that correspond to consistency on GSI score levels over time, e.g. the largest cells show medium scores for all four waves.

We are interested in modeling the log odds-ratios or logits for GSI score at wave 4, e.g.,

$$\log (m_{ijk1d}/m_{ijk2d}) \text{ and } \log (m_{ijk2d}/m_{ijk3d}) \quad (2)$$

as a function of the four explanatory variables. Actually, as noted in Fienberg (2) it is somewhat more convenient to model the log continuation ratios:

$$\log [m_{ijk1d}/(m_{ijk2d} + m_{ijk3d})] \text{ and } \log (m_{ijk2d}/m_{ijk3d}) \quad (3)$$

Table 2. TMI data for GSI scores.

Distance	Wave 1	Wave 2	Wave 3	Wave 4		
				Low4	Med4	High4
LTE 5	Low1	Low2	Low3	2	0	0
			Med3	2	3	0
			High3	0	0	0
		Med2	Low3	0	1	0
			Med3	2	4	0
			High3	0	0	0
		High2	Low3	0	0	0
			Med3	0	0	0
			High3	0	0	0
		Med1	Low2	Low3	1	0
				Med3	4	0
				High3	0	0
			Med2	Low3	2	0
				Med3	38	4
				High3	2	3
			High2	Low3	0	0
				Med3	2	0
				High3	1	1
	High1	Low2	Low3	0	0	0
			Med3	0	0	0
			High3	0	0	0
		Med2	Low3	0	0	0
			Med3	0	4	3
			High3	0	1	4
		High2	Low3	0	0	0
			Med3	1	2	0
			High3	0	5	12
GT5	Low1	Low2	Low3	1	2	0
			Med3	2	0	0
			High3	0	0	0
		Med2	Low3	1	0	0
			Med3	0	3	0
			High3	0	0	0
		High2	Low3	0	0	0
			Med3	0	0	0
			High3	0	0	0
		Med1	Low2	Low3	4	0
				Med3	15	1
				High3	0	0
			Med2	Low3	2	0
				Med3	53	6
				High3	5	1
		High2	Low3	0	0	0
			Med3	0	1	1
			High3	0	3	1
	High1	Low2	Low3	0	0	1
			Med3	0	0	0
			High3	0	0	0
		Med2	Low3	0	0	0
			Med3	1	13	0
			High3	0	0	0
		High2	Low3	0	0	0
			Med3	0	7	2
			High3	0	2	7

as the likelihood function factors into separate likelihoods for the two ratios. For example, we might consider fitting a simple additive model, to each log ratio, of the form

$$w + w_{1(i)} + w_{2(j)} + w_{3(k)} + w_{5(d)} \quad (4)$$

where each set of w -terms adds to zero when summed over the subscript. These models can be fit using the method of iterative proportional fitting (as in the BMDP 4F program) or using a version of the Newton-Raphson method, as in the GLIM program of Baker and Nelder (20).

A summary of the fit of selected versions of model (4) is given in Table 3. The fit of the additive model appears to be good with each of the G^2 statistics (minus twice the log likelihood ratio) being roughly equal to the corresponding degrees of freedom (df). The second model listed in Table 3 drops the effect of distance, and the differences in G^2 values,

$$23.0 - 22.7 = 0.3$$

and

$$26.9 - 25.5 = 1.4$$

$$\Delta df = 1$$

indicate that distance as a variable has negligible predictive power.

Finally the last model in the table drops the effects of GSI score at waves 1 and 2. This model has a Markov-chainlike interpretation—GSI at wave 4 is independent of GSI at waves 1 and 2 given GSI at wave 3. If we take differences in loglikelihood ratio statistics for models (ii) and (iii) we get

$$\Delta G_1^2 = 35.6 - 23.0 = 12.6$$

$$\Delta df = 4$$

$$\Delta G_2^2 = 36 - 26.9 = 9.1$$

When we refer these differences to a χ^2 distribution on 4 df, the values are significant at the 0.025 and 0.10 levels, respectively. Thus there appears to be some residual effects of psychological symptomatology from waves 1 and 2 on the GSI scores at wave 4, even after adjusting for the GSI level at wave 3. These and other related logit model analyses are described in detail in May (21).

At this stage we reintroduce the information on perception of danger, as measured at each wave, and we consider perception as a risk factor for psychological stress. The $3^4 \times 2$ table we have been working with is a marginal table for the full $(3 \times 2)^4 \times 2$ array. Separate analyses of the 3×2 tables for GSI score by perception for each of the four waves show that perception of danger is associated to a modest degree with higher levels of GSI. Finally an examination of the 2^4 table giving perception of danger for the four waves suggests that mothers tend to give the same answer to the question of perception of danger from one wave to the next. We are interested in developing an overall model for the data which attempts to incorporate these preliminary results from the various marginal arrays.

We can combine the information from the various preliminary analyses in diagrammatic form as follows:

$$\begin{array}{ccccccc} G_1 & \rightarrow & G_2 & \rightarrow & G_3 & \rightarrow & G_4 \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ P_1 & \rightarrow & P_2 & \rightarrow & P_3 & \rightarrow & P_4 \end{array} \quad (5)$$

where P_i and G_i represent perception of danger and GSI scores, respectively, in wave i for $i = 1, 2, 3, 4$. The horizontal arrows suggest that there is a Markovian-like form of persistence for both GSI symptomatology and perception of danger over time. The vertical arrays suggest that perception of danger continues to reinforce the existence of psychological stress. For the moment, we ignore the distance variable.

We can formulate a loglinear model for the $(3 \times 2)^4$ table that includes the various relationships depicted in the figure of expression (5) or we can re-express that model as a sequence of logit four models, treating GSI at wave i for $i = 1, 2, 3, 4$ as a response. For the i -th model the explanatory variables are perception of danger at wave i and GSI and perception at all previous waves. Within this sequence of models we can test, for waves 2, 3, and 4, whether perception of danger appears to reinforce psychological stress, i.e., we can test whether the parameters linking P_i to G_i are equal to zero. Although we omit the details of the analysis of these loglinear and logit models here, they tend to provide little support for the effect of perception of danger on GSI scores once we adjust for prior GSI scores. The loglinear and logit model analyses just described are

Table 3. The fit of selected versions of the additive logit model of expression (3).

Model	df	Log likelihood ratio chi-square statistic (G^2)	
		$\log \left(\frac{m_{ijk1d}}{m_{ijk2d} + m_{ijk3d}} \right)$	$\log \left(\frac{m_{ijk2d}}{m_{ijk3d}} \right)$
(i) $w + w_{1(i)} + w_{2(j)} + w_{3(k)} + w_{5(d)}$	22	22.7	25.5
(ii) $w + w_{1(i)} + w_{2(j)} + w_{3(k)}$	23	23.0	26.9
(iii) $w + w_{3(k)}$	27	35.6	36.0

Table 4.

Cell	Observed	Expected
000	159	129.0
001	19	30.5
010	14	28.5
011	11	13.1
100	16	30.6
101	7	11.5
110	12	12.9
111	29	10.3

consistent with a set of regression analyses in which the logarithms of the original GSI scores are regressed on current perception and prior log-GSI values.

On the basis of these analyses, we can only conclude that perception of danger is not a substantial risk factor for psychological distress. Before we rush to do so, we need to reexamine the data. We know that those with a consistently low GSI score tend to cluster at the two extremes in the response to the question on perception of danger—consistent responses of “yes” or consistent responses of “no”. Thus, there may well be a latent variable corresponding to “background propensity” that we need to take into account before attempting to assess the impact of perception on GSI level. In the next section, we describe a logistic response model that incorporates a structure allowing for such individual differences.

A Mixture Approach to Logit Models

For purposes of simplicity, suppose that we are interested in a dichotomous response variable and that we are working with a logistic response model of the form

$$\log \frac{p}{1-p} = a + \sum \beta_i x_i \quad (6)$$

This is the same model as in expression (1) but choose to highlight the “intercept” term, a . This model assumes that the population of all individuals with the same values for the explanatory variables, $\{x_i\}$, is homogeneous. Suppose, however, that there is a heterogeneity that is reflected in differential propensities to yield a positive response beyond that accounted for by the explanatory variables. This heterogeneity can be modelled by letting a be a random variable with distribution function H . Then, if the observed response has a Bernoulli distribution, the likelihood associated with a response y (which is 0 or 1) is

$$l(H, \beta \mid y, x) = \left(\int \left[\frac{e^{a + \sum \beta_i x_i}}{1 + e^{a + \sum \beta_i x_i}} \right] dH(a) \right)^y \left(\int \left[\frac{1}{1 + e^{a + \sum \beta_i x_i}} \right] dH(a) \right)^{1-y} \quad (7)$$

Follmann (22) studies the properties of these mixture models and considers the estimation of the β_i when the

mixing distribution, H , is unknown. For other approaches to the problem of extra-binomial variation see Williams (23) and the references therein.

We now consider an extension to this basic mixture model that is suitable to the kind of situation depicted in the figure of expression (5) above. Let Y_i be the random variable giving a dichotomized GSI score at wave i and let \mathbf{x}_i be a vector of covariates measured at time i . Then we let

$$P(Y_i = 1 \mid \mathbf{x}_i, y_{i-1}) = \frac{e^{a + \beta' \mathbf{x}_i + \gamma y_{i-1}}}{1 + e^{a + \beta' \mathbf{x}_i + \gamma y_{i-1}}} \quad (8)$$

$$= P(Y_i \mid a) \quad i = 2, 3, 4$$

where y_{i-1} is the observed response in the previous wave. The contribution for an individual to the likelihood comes from putting together the Bernoulli contribution for waves 2, 3, and 4, then giving a a distribution $H(a)$, and finally integrating over H , i.e.,

$$P(Y_2 = y_2, Y_3 = y_3, Y_4 = y_4) = \int P(Y_2 \mid a)^{y_2} [1 - P(Y_2 \mid a)]^{1-y_2} P(Y_3 \mid a)^{y_3} [1 - P(Y_3 \mid a)]^{1-y_3} P(Y_4 \mid a)^{y_4} [1 - P(Y_4 \mid a)]^{1-y_4} dH(a) \quad (9)$$

The likelihood is found by taking a product of these probabilities over all individuals in the sample. We estimate H , γ , and β via the method of maximum likelihood.

The mixture model when considered as an alternative to an unmixed model can lead to insight in two major ways. First, the mixing distribution, H , may be interpretable. Second, the pattern between GSI, time, and preception of danger may be subjected to so much ‘noise’ that spurious conclusions may be drawn unless the noise is accommodated.

To ensure identifiability, we require that H be a discrete distribution with a bound C on the number of support points. That is, there are a finite number of possible values for a (a_1, a_2, \dots, a_c) with associated probabilities q_1, q_2, \dots, q_c where $c \leq C$. This bound depends on the observed covariates (21). Since the number of support points c is unknown, this leads to some tricky theoretical problems when we attempt to check the goodness of fit of our model given by Eqs. (8) and (9), but we do not dwell on such problems here.

The actual computation of parameter estimates requires a different computing algorithm than those that are used for logistic response models. Follmann (21) describes an adaptation of an algorithm described by Everitt and Hand (24) which can be viewed as a special case of the EM algorithm of Dempster, Laird, and Rubin (25). For the TMI analyses described in this paper, we have an estimated mixing distribution, \hat{H} with $\hat{c} = 2$ components. The EM algorithm as implemented here is very slow to converge—especially when contrasted with the iterative scaling or GLIM algorithms referred to above.

Mixture Model Analysis of TMI Data

We now apply the mixture model to the TMI data. Our purpose is to explore where adjusting for the differing background propensities to view TMI as dangerous changes the overall conclusions from the analyses of the previous section.

As explanatory variables in the logistic response structure of expression (8) we have: (i) current perception of danger (CP), (ii) distance from TMI (DIST), (iii) dichotomized GSI at previous wave (GSIL). For all of the variants on this model which we examined, although we allowed for a mixing distribution with 2 or more values for a , from the data we estimate that there is only one group, i.e., $\hat{c} = 1$. We tried various cutpoints for dichotomizing GSI scores and they yielded basically similar results. For the following results we used a wave-dependent cutpoint, which classified the lower 3/4 of the ranked GSI scores as 0 and the upper 1/4 as 1:

logit (GSI) =

$$-1.65 + 0.39CP - 0.95DIST + 2.41 GSIL$$

(−5.96) (1.78) (−2.82) (12.36)

The numbers in parentheses are coefficients divided by their asymptotic standard deviations. For the key coefficient of interest, corresponding to the effect of perception of danger on the odds of high GSI score, the standardized value of 1.78 is marginally significant (between $p = 0.10$ and $p = 0.05$). Note that distance from TMI which is included here directly (not in a dichotomized form), does have a significant effect. To illustrate how well this model fits the data we have estimated expected cell values for the marginal 2^3 table for GSI scores at waves 2, 3, and 4 (Table 4). The fit of the model is clearly poor.

We can also consider a model which reverses the vertical arrows in the figure of expression (3), i.e., consider a model which predicts perception of danger from GSI level. As explanatory variables we have (i) current GSI (CGSI), (ii) distance (DIST), (iii) perception at previous wave (PL), and (iv) a dummy variable which is 0 for wave 2 and 1 for waves 3 and 4 (WAVE). The GSI scores here do not necessarily need to be dichotomized and in the analyses reported below they are untransformed. Here we fit both the standard logistic response model and the mixture model. In the latter case the estimated

number of groups is $\hat{c} = 2$, so that we get two estimated intercept terms (Table 5). For the mixture model, the standard errors were computed assuming that c was known to be 2 a priori. These are two clearly identified components in our analysis and this leads to our estimate of $\hat{c} = 2$. It remains possible that $c > 2$, and thus our assumption of a known value of c is likely to make the standard errors somewhat smaller than they should be, and the standardized coefficients a little bit too large.

The estimated mixing distribution is

$$\hat{H} = 0.39\delta_{-1.93} + 0.61\delta_{1.55}$$

The distribution puts probability 0.39 on $\hat{a}_1 = -1.93$ and probability 0.61 on $\hat{a}_2 = 1.55$.

In both analyses there now seems to be a significant link between GSI scores and perception of danger. The two analyses can be contrasted in several different ways. First, the difference in the regression coefficients for the covariates are minimal. The main difference is a reduction in the magnitude of the coefficient for perception of danger of the previous wave. Second, the mixture model fits much better than the unmixed model. The two models differ by only two parameters and $-2 [\text{loglikelihood (NO-MIX)} - \text{loglikelihood (MIX)}] = 2 (431.46 - 403.44) = 56.04$. Referring this value to a χ^2_2 reference distribution, we see that it is highly significant, i.e., there is an important improvement in fit associated with the mixture model. Under a nonmixture model, we are assuming that the only difference in the probability of an outcome lies in differences in the covariate values. Thus, if two mothers have the same covariate values, they have the same probability of the outcome. The mixture model permits two mothers with the same covariate values to have two different probabilities of the outcome, corresponding to high and low propensities to perceive the reactor as dangerous.

As an indicator of the fit of the mixture models we have computed estimated expected values for both the mixture and nonmixture models for the 2^3 marginal table of perception of danger at waves 2, 3, and 4 (Table 6). The improvement in the fit of the mixture model over the nonmixture model is reflected in these estimated expected values. The major residual discrepancy occurs in the 000 cell, but it is not excessively large.

The interpretation of the key estimated coefficients

Table 5.

	No mixing		Mixing	
	Coefficient	Standardized	Coefficient	Standardized
Intercept	− 0.676	− 2.401	− 1.926	− 3.761
			1.554	2.988
PL	1.685	9.775	0.641	2.387
CGSI	1.003	3.428	1.260	3.046
DIST	.042	1.433	.096	1.794
WAVE 2–0	− 0.294	− 1.637	− 0.404	− 1.831
Loglikelihood	− 431.46		− 403.44	

Table 6.

Cell	Observed	Unmixed expected	Mixed expected
000	39	9.4	30.5
001	8	13.7	12.9
010	9	15.2	13.4
011	15	32.5	13.1
100	25	22.7	19.7
101	13	37.0	18.7
110	14	41.4	19.8
111	144	94.6	138.0

within the mixture model is as follows. The odds for the probability of the perception of danger are increased by (i) perception of danger at the previous wave, (ii) larger current GSI scores and (iii) increased distance from TMI. (This latter variable has a different sign than we expected to find but its standardized coefficient is only marginally significant.)

Additional analyses involving this mixture model that allow for different coefficients at different waves do not appreciably change either the fit of model or the general interpretation of the coefficients. We are left with somewhat mixed evidence on the implications of the perception of danger as a risk factor for psychological stress in the aftermath of the TMI accident. Yet enough evidence has emerged in our analyses to warrant close scrutiny of this aspect of psychological stress in connection with other large-scale environmental hazards with potentially life threatening implications.

This research was supported by NIMH Contract No. 278-79-0048, and by an NIMH Training grant, No. 2 T32 MH15758-06. This paper was prepared for presentation at Environment Risk Assessment and Statistical Methods Conference, Kyoto, Japan, August 30-31, 1984, while the first author was supported by a Guggenheim fellowship. We thank C. Ahn, C. Kim, K. Morrissey, and E. Sungur for their work on some of the preliminary analyses.

REFERENCES

1. Breslow, N. Cohort analysis in epidemiology. In: *A Celebration of Statistics: The ISI Centenary Volume* (A. C. Atkinson and S. E. Fienberg, Eds.), Springer Verlag, 1985, pp. 107-143.
2. Fienberg, S. E. *The Analysis of Cross-classified Categorical Data*. (2nd ed.). Massachusetts Institute of Technology Press, Cambridge, 1980.
3. Fienberg, S. E. Towards a comprehensive approach to the analysis of categorical data. In: *Statistics: An Appraisal* (H. A. David and H. T. David, Eds.), Iowa State University Press, Ames, 1984, pp. 393-422.
4. Imrey, P. B., Koch, G. G., Stokes, M. E., Darroch, J. N., Freedman, D. H., Jr., and Tolley, H. D. Categorical data analysis: some reflections on the loglinear model and logistic regression. Part I: historical and methodological overview. *Int. Statist. Rev.* 49: 265-284 (1981).
5. Yule, G. U. On the association of attributes in statistics: with illustration from the material of the childhood society. *Phil. Trans. Roy. Soc. A* 194: 257-319 (1900).
6. Bartlett, M. S. Contingency table interactions. *J. Roy. Statist. Soc. (Suppl.)* 2: 248-252 (1935).
7. Dyke, G. V., and Patterson, H. D. Analysis of factorial arrangements when the data are proportions. *Biometrics* 8: 1-12 (1952).
8. Dawber, T. R. *The Framingham Study: The Epidemiology of Atherosclerotic Disease*. Harvard University Press, Cambridge, 1980.
9. Bunker, J. P., Forrest, W. H., Jr., Mosteller, F., and Vandam, L. *The National Halothane Study. Report of the Subcommittee on the National Halothane Study of the Committee on Anesthesia, Division of Medical Sciences, National Academy of Sciences-National Research Council, National Institutes of Health, National Institute of General Medical Sciences, Bethesda, MD, U.S. Government Printing Office Washington, DC, 1969.*
10. Truett, J., Cornfield, J., and Kannel, W. A multivariate analysis of the risk of coronary heart disease in Framingham. *J. Chron. Dis.* 20: 511-524 (1967).
11. Walker, S. H., and Duncan, D. B. Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54: 167-179 (1967).
12. Bishop, Y. M. M. *Multidimensional contingency tables; cell estimates*. Ph.D. Dissertation. Department of Statistics, Harvard University, 1967; available from University Microfilm Service, Ann Arbor, MI.
13. Deming, W. E., and Stephan, F. F. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Statist.* 11: 427-444 (1940).
14. Birch, M. W. Maximum likelihood in three-way contingency tables. *J. Roy. Statist. Soc. B* 25: 220-233 (1963).
15. Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. *Discrete Multivariate Analysis: Theory and Practice*. Massachusetts Institute of Technology Press, Cambridge, 1975.
16. Bromet, E., Parkinson, D., Schulberg, H., Dunn, L., and Gondek, P. Mental health of residents near the Three Mile Island reactor: a comparative study of selected groups. *J. Prev. Psychiat.* 1: 225-276 (1982).
17. Moss, P., and Plewis, I. Mental distress in mothers of pre-school children in Inner London. *Psychol. Med.* 7: 641-652 (1977).
18. Derogotis, L. R. *The SCL-90 Manual I: Scoring, administration and procedures for the SCL-90*. The Johns Hopkins University School of Medicine, Clinical Psychometrics Unit, Baltimore, 1977.
19. Grizzle, J. E., and Koch, G. G. Some applications of categorical data analysis to epidemiological studies. *Environ. Health Perspect.* 32: 169-179 (1979).
20. Baker, R. J., and Nelder, J. A. *The GLIM System Manual (Release 3)*. The Numerical Algorithms Group/Royal Statistical Society Oxford, 1978.
21. May, S. M. Log-linear modelling techniques applied to some epidemiologic and psychiatric data. Ph.D. Dissertation, Department of Statistics, Carnegie-Mellon University, 1984.
22. Follmann, D. Nonparametric mixtures in logistic regression. Ph.D. Dissertation, Department of Statistics, Carnegie-Mellon University, 1985.
23. Williams, D. A. Extra-binomial variation in logistic linear models. *Appl. Statist.* 31: 144-148 (1982).
24. Everitt, B. S., and Hand, D. J. *Finite Mixture Distributions*. Chapman and Hall, New York, 1981.
25. Dempster, A. P., Laird, N., and Rubin, D. B. Maximum likelihood estimation with incomplete data via the EM algorithm. *J. Roy. Statist. Soc. B* 39: 1-38 (1977).